

This local analysis of melioration is largely consistent with the position of Donahoe et al. Like theirs, this approach treats strength of responding as part of the three-term contingency. On the other hand, this analysis assumes two equations for learning, one for classical and one for operant behavior. Perhaps Equation 5 could be generalized to cover classical conditioning, in which case we would be back to a single process, but the equation would represent a break from the Rescorla–Wagner model.

REFERENCES

- Baum, W. M. (1973). The correlation-based law of effect. *Journal of the Experimental Analysis of Behavior*, 20, 137–153.
- Dinsmoor, J. A. (1962). Variable-interval escape from stimuli accompanied by shocks. *Journal of the Experimental Analysis of Behavior*, 5, 41–47.
- Herrnstein, R. J., & Vaughan, W., Jr. (1980). Melioration and behavioral allocation. In J. E. R. Staddon (Ed.), *Limits to action* (pp. 143–176). New York: Academic Press.
- Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning: II. Current research and theory* (pp. 64–99). New York: Appleton-Century-Crofts.
- Vaughan, W., Jr. (1982). Choice and the Rescorla–Wagner model. In M. L. Commons, R. J. Herrnstein, & H. Rachlin (Eds.), *Quantitative analyses of behavior: Vol. 2. Matching and maximizing accounts* (pp. 263–279). Cambridge, MA: Ballinger.
- Williams, B. A. (1976). The effects of unsignalled delay of reinforcement. *Journal of the Experimental Analysis of Behavior*, 26, 441–449.

WHAT IS LEARNED?
REVISITING AN OLD ISSUE

BEN A. WILLIAMS

UNIVERSITY OF CALIFORNIA, SAN DIEGO

The authors of this provocative article argue that an S-R approach to behavior is an implicit assumption of connectionist network models of behavior. More exactly, their S-R model is better depicted as an S-O-R model, because a large part of variability in behavior comes not from variation in the stimulus input but from differences in activity, including spontaneous activity, between the nodes of the intervening network. As the authors make clear, these assumptions make it difficult to distinguish their S-R account from the traditional operant analysis in terms of the three-term contingency.

Although the authors make clear that their

type of S-R analysis does not necessarily require a causal explanation of behavior in terms of the particular stimulus that elicits each response, their approach does share with other S-R approaches (e.g., Thorndike, Hull) the assumption that the function of the reinforcer is to provide catalysis of S-R associations but not to enter into the associative relation itself. Donahoe et al.'s model utilizes the release of dopamine as the agent that increases the connection weights between different elements of the network. This so-called reinforcer, like the reinforcer in traditional S-R theory, remains outside of the associative network.

A critical issue posed for Donahoe et al. is whether a satisfactory model of conditioning can be constructed that omits any role for response–reinforcer associative relations. Al-

Correspondence concerning this article should be addressed to Ben A. Williams, Psychology Department, University of California San Diego, La Jolla, California 92093-0109 (E-mail: bawilliams@ucsd.edu).

though Skinner himself eschewed any such analysis, other prominent behavior analysts have questioned whether the role of the reinforcer can be understood simply in terms of catalysis of the S-R connections. For example, Catania (1984, p. 77) writes, "The consequences of responding are critical to our understanding of learning not because learning follows from them but because they are what is learned." Subsequent empirical developments in associative learning theory have strongly supported Catania's perspective that the response-reinforcer relation is an essential part of what is learned in operant behavior. Much of this work comes from studies of reinforcer devaluation. For example, Colwill and Rescorla (1986) trained separate responses using different reinforcers for each response, and then independently (with the responses no longer available) devalued one of the reinforcers by pairing it with poison. The subject was then returned to the training situation with both responses freely available but during extinction. The results were that only the response that led to the devalued reinforcer in the past was decreased in strength. The specificity of the effect seems most easily interpretable in terms of specific response-reinforcer associations. Donahoe et al. are clearly aware of the importance of these devaluation studies, in that Donahoe and Palmer (1994, pp. 108-109) provide an explanation of the basic effect. However, their treatment of the issue commits them to an interpretation in terms of two-factor theory, in which the animal's conditioned anticipation of the reward produces stimuli that assume discriminative control. Although such an interpretation has not been totally ruled out, variations in the procedure (see Colwill & Rescorla, 1986) that cause the same external stimulus to be present for both responses (e.g., with a bidirectional lever, with different reinforcers contingent on the different directions of the response) do seem to cause the two-factor perspective to be strained in its application.

Perhaps the most direct evidence that the discriminative stimulus serves to cue the response-reinforcer relation in effect, and thus cannot be reduced to being a direct elicitor of responding, comes from studies of stimulus control that use blocking as a tool for analysis (Colwill & Rescorla, 1990, Experi-

ment 3; Rescorla, 1990). In the pretraining phase of these studies, each of two responses led to one of two different reinforcers in the presence of a discriminative stimulus (A), and neither reinforcer was available during the absence of the stimulus. In the compound conditioning phase, two new stimuli were added to form two separate stimulus compounds, BA and CA. In the presence of BA, the response-reinforcer relations were the same as those used during pretraining, whereas in the presence of CA, the opposite response-reinforcer relations were in effect. Then Elements B and C were tested alone to determine if they had acquired stimulus control over responding. In both experiments, Element B failed to acquire stimulus control, indicating blocking, but Element C did acquire control. The most plausible interpretation of this pattern of results is that stimulus control is determined by whether the stimulus provides new information about the response-reinforcer relation. The challenge I would pose for Donahoe et al. is to demonstrate that their connectionist model can account for this array of results. Given that the reinforcer itself is not represented in the network, it is not obvious how such an account would proceed.

The issue of what terms enter into associative relations is also important to the authors' distinction between Pavlovian and instrumental conditioning. They are certainly correct that responses necessarily are occurring in the presence of stimuli at the time of reinforcement and that the major difference between the two kinds of contingencies is the degree of constraint on the co-occurrence of the different events. But a large amount of evidence argues that the associative relation in Pavlovian conditioning cannot be reduced to S-R connections. For example, if a tone is paired with food in an autoshaping procedure, it elicits orientation but no pecking. If the tone is then made contingent on the prior occurrence of a keylight, pecking to the keylight will occur. What appears to have occurred is that the tone has assumed the status of a substitute for food, and the behavior it elicits is some combination of the unconditioned effects of the reinforcer and of the particular stimulus that serves as the CS. A major assumption of S-R theory is that the behavior to a stimulus is what is learned, so

that the interpretation of Pavlovian conditioning in terms of sheer stimulus substitution omits a critical ingredient of the S-R associative relation. The evidence for a stimulus substitution interpretation, instead of an S-R association, is, to this reader, compelling (see Mackintosh, 1983, for a review and analysis).

Quite apart from the issue of what is learned, there are several other issues that prompt Donahoe et al. to endorse a theoretical position that is open to challenge. Whether these positions are intrinsic to their connectionist model is unclear. The most general is their commitment to a moment-by-moment analysis of phenomena such as the matching law in terms of changing stimulus states, even when the available evidence seems to support the molar interpretation. Because the molar-molecular issue has been debated extensively in the literature, it may be useful to consider some of its implications for their connectionist model.

A critical issue is the nature of the choice rule that governs response selection. S-R analyses implicitly assume that the rule is maximizing, or winner-take-all, in that whichever response has the greatest associative connection, given the stimulus set, will be evoked. In contrast is the matching rule, which assumes that the different responses are emitted proportionally to their strength. The issue is not yet resolved (see Williams, 1994, for a discussion), but Donahoe et al. argue that the evidence favors the momentary maximizing approach when the two different perspectives are pitted against each other. I disagree with this interpretation of the evidence. Although it is true that momentary contingencies, defined by fluctuating stimulus states, can control behavior when they are sufficiently strong, it does not appear to be the case that the molar phenomena that have dominated operant research over the past 30 years are easily reduced to the molecular effects. The results of Williams (1991) provide an example. In that study rats were trained on a discrete-trial probability learning task in which the local contingencies always favored win-shift and lose-shift response patterns (i.e., alternation), but were asymmetric for the two responses, causing the molar reinforcement rates for the two responses to differ. The results were that some modicum of control by the local contingencies did occur with short

intertrial intervals, but these contingencies were discriminated very poorly with longer intertrial intervals. Most important, approximation to the matching law was much closer with the longer ITIs, when the local contingencies had little control over behavior. The issue posed for Donahoe et al. is how to provide a moment-by-moment analysis of behavior, in this case choice behavior, when the contingencies associated with the local stimulus environment apparently cannot account for the obtained pattern of behavior.

An especially challenging example of the inadequacy of molecular models is provided by Neuringer (1992). Rats' responding was reinforced on separate VI schedules when that behavior met a criterion for repetition or for randomness. The result was approximate matching of the two types of behavior to their relative reinforcement frequencies. Given that one of the response alternatives could not be predicted on a moment-by-moment basis, and that prior work (Page & Neuringer, 1985) had made a strong case that learning to emit random behavior could not be the result of a memory-based strategy, the functional equivalence of such behavior with normal behavior with known controlling stimuli poses a major conceptual challenge for any type of S-R analysis.

A final issue that deserves consideration is the status of associationism in general. Gallistel (1990) has provided a strong cogent argument that learning does not consist of the formation of associative links, but instead consists of representations of the structure of the environment. In other words, the animal acquires veridical knowledge of the spatial and temporal properties of important events, which is mediated only indirectly by how those events make contact with behavior via a reinforcement contingency. For example, in his view, the matching law is the result of the animal having knowledge of the different distributions of interreinforcement intervals contingent on each response alternative, and is not the result of the number of pairings of the response and reinforcer, or of the probability that a response will be followed by a reinforcer. Substantial evidence (Gibbon, Church, Fairhurst, & Kacelnik, 1988; Mark & Gallistel, 1994; Williams, 1993) now exists in favor of such an analysis. Obviously, it is premature to pass judgment on such an impor-

tant overarching issue. Whether associative models can deal with the evidence compiled by Gallistel (1990) remains an open question, and only direct simulations of those findings will provide an answer. It is my hope that Donahoe et al. will take up this challenge and the others previously described. Connectionist networks will rise or fall as psychological explanations to the extent that they can simulate the properties of real behavioral events.

REFERENCES

- Catania, A. C. (1984). *Learning* (2nd ed.). Englewood Cliffs, NJ: Prentice Hall.
- Colwill, R. M., & Rescorla, R. A. (1986). Associative structures in instrumental learning. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 20, pp. 55–104). New York: Academic Press.
- Colwill, R. M., & Rescorla, R. A. (1990). Evidence for the hierarchical structure of instrumental learning. *Animal Learning & Behavior*, 18, 71–82.
- Donahoe, J. W., & Palmer, D. C. (1994). *Learning and complex behavior*. Boston: Allyn and Bacon.
- Gallistel, C. R. (1990). *The organization of learning*. Cambridge, MA: MIT Press.
- Gibbon, J., Church, R. M., Fairhurst, S., & Kacelnik, A. (1988). Scalar expectancy theory and choice between delayed rewards. *Psychological Review*, 95, 102–114.
- Mackintosh, N. J. (1983). *Conditioning and associative learning*. Oxford: Oxford University Press.
- Mark, T. A., & Gallistel, C. R. (1994). Kinetics of matching. *Journal of Experimental Psychology: Animal Behavior Processes*, 20, 79–95.
- Neuringer, A. (1992). Choosing to vary and repeat. *Psychological Science*, 3, 246–250.
- Page, S., & Neuringer, A. (1985). Variability is an operant. *Journal of Experimental Psychology: Animal Behavior Processes*, 11, 429–452.
- Rescorla, R. A. (1990). The role of information about the response-outcome relation in instrumental discrimination learning. *Journal of Experimental Psychology: Animal Behavior Processes*, 16, 262–270.
- Williams, B. A. (1991). Choice as a function of local versus molar reinforcement contingencies. *Journal of the Experimental Analysis of Behavior*, 56, 455–473.
- Williams, B. A. (1993). Molar versus local reinforcement probability as determinants of stimulus value. *Journal of the Experimental Analysis of Behavior*, 59, 163–172.
- Williams, B. A. (1994). Reinforcement and choice. In N. J. Mackintosh (Ed.), *Animal learning and cognition* (pp. 81–108). San Diego: Academic Press.